

Scaling and memory in volatility return intervals in financial markets

Kazuko Yamasaki^{*†}, Lev Muchnik[‡], Shlomo Havlin^{*§}, Armin Bunde[¶], and H. Eugene Stanley^{*}

^{*}Center for Polymer Studies and Department of Physics, Boston University, Boston, MA 02215; [†]Department of Environmental Sciences, Tokyo University of Information Sciences, Chiba 265-8501, Japan; [‡]Minerva Center and Department of Physics, Bar-Ilan University, Ramat-Gan 52900, Israel; and [¶]Institut für Theoretische Physik III, Justus-Liebig-Universität, D-35392 Giessen, Germany

Contributed by H. Eugene Stanley, April 1, 2005

For both stock and currency markets, we study the return intervals τ between the daily volatilities of the price changes that are above a certain threshold q . We find that the distribution function $P_q(\tau)$ scales with the mean return interval $\bar{\tau}$ as $P_q(\tau) = \bar{\tau}^{-1} f(\tau/\bar{\tau})$. The scaling function $f(x)$ is similar in form for all seven stocks and for all seven currency databases analyzed, and $f(x)$ is consistent with a power-law form, $f(x) \sim x^{-\gamma}$ with $\gamma \approx 2$. We also quantify how the conditional distribution $P_q(\tau|\tau_0)$ depends on the previous return interval τ_0 and find that small (or large) return intervals are more likely to be followed by small (or large) return intervals. This “clustering” of the volatility return intervals is a previously unrecognized phenomenon that we relate to the long-term correlations known to be present in the volatility.

econophysics | fluctuations | extreme values | long-term correlations | long-term memory

The statistical properties of stock and currency market fluctuations are of importance for modeling and understanding complex market dynamics. They are also relevant for practical applications such as risk estimation and portfolio optimization (1). In particular, understanding the volatility fluctuations of financial records is of particular importance, because they are the key input of option pricing models, including the classic Black and Scholes model and the Cox, Ross, and Rubinstein binomial models that are based on estimates of the asset's volatility during the residual time of the option (2–4). Although the changes from day $i - 1$ to day i , $\Delta p_i \equiv p_i - p_{i-1}$, of both stock prices and currency rates are uncorrelated, their absolute values (one measure of volatility) are long-term power-law correlated (5–17). Moreover, the probability density function (pdf) of Δp_i scales as a power law (18) $\Phi(\Delta p) \sim (\Delta p)^{-(\zeta + 1)}$ with $\zeta \approx 3$ (5, 19–21). Also, within t days after a crash, $n_q(t)$, the number of times $|\Delta p_i|$ exceeds a threshold q , follows a power-law relation $n_q(t) \sim t^{-p}$ with $p \approx 1$ (22), a behavior similar to the Omori earthquake law.

Here, we are interested in the statistical properties of large volatilities. A quantity that characterizes the occurrence of large volatilities is the return interval τ between two consecutive volatilities above some large threshold q (Fig. 1). We study return intervals because they are related to the rate of occurrence of volatilities that exceed a threshold q (22). Because extreme volatilities are rare, we consider also the return intervals between volatilities above intermediate thresholds. By doing this, we hope to gain insight also into the return intervals between very large volatilities that are too rare to obtain with reasonable statistics.

We analyze the statistical properties of the daily return intervals of seven representative stocks and currencies obtained, respectively, from <http://finance.yahoo.com> and www.federalreserve.gov/releases/H10/hist. We choose to study daily data records because there are intraday trends in the volatility. We report two results:

(i) The pdf of the return intervals $P_q(\tau)$ is not a function of the two independent variables τ and q but depends only on the scaled parameter $\tau/\bar{\tau}$, where the q dependence is contained in the mean

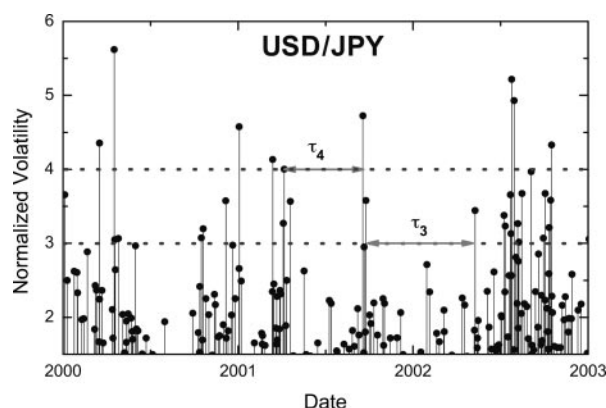


Fig. 1. Schematic illustration of volatility return intervals. Shown are the return intervals τ_3 and τ_4 for two threshold values $q = 3$ and 4 (indicated by arrows) for the normalized volatility $|\Delta p_i|/((\langle |\Delta p_i|^2 \rangle - \langle \Delta p_i \rangle^2)^{1/2})$ of USD/JPY currency exchange rates in the 3-year period 2000–2002.

return interval $\bar{\tau} \equiv \bar{\tau}(q)$. This scaling is important because it allows us to extrapolate large q values, corresponding to rare events, from the behavior at small q values and thereby collect good statistics. We find that the scaled pdf can be well approximated by the same power law for all seven stocks and for all seven currencies studied. We also show that this power-law behavior results from the known long-term correlations in the volatility records.

(ii) A long-term memory exists in the return intervals, such that short return intervals tend to be followed by short ones, and long return intervals tend to be followed by long ones. This “clustering” of the volatility return intervals is a previously unrecognized phenomenon that we relate to the long-term correlations known to be present in the volatility.

Methods

We normalize the volatility records by dividing each value of $|\Delta p_i|$ by the standard deviation $((\langle |\Delta p_i|^2 \rangle - \langle \Delta p_i \rangle^2)^{1/2})$. In this way, the thresholds q are in units of the standard deviation of the volatility. Because we analyze daily records, we restrict ourselves to those q values for which $\bar{\tau} > 3$ days to avoid spurious discreteness effects.

We begin by studying the behavior of the pdf $P_q(\tau)$ and how it depends on the threshold parameter q (Fig. 2 *a* and *b*). For different q , the pdfs are different and cannot be described by a Poisson distribution as for uncorrelated data. Moreover, for larger q , the decay of $P_q(\tau)$ becomes slower.

To understand the q dependence, we show, in Fig. 2 *c* and *d*, the scaled pdfs $P_q(\tau)\bar{\tau}$ as functions of the scaled return intervals

Abbreviations: pdf, probability density function; S&P 500, Standard and Poor's 500 Index; USD, U.S. dollar; JPY, Japanese yen; SEK, Swedish krona.

[§]To whom correspondence should be addressed. E-mail: havlin@ophir.ph.biv.ac.il.

© 2005 by The National Academy of Sciences of the USA

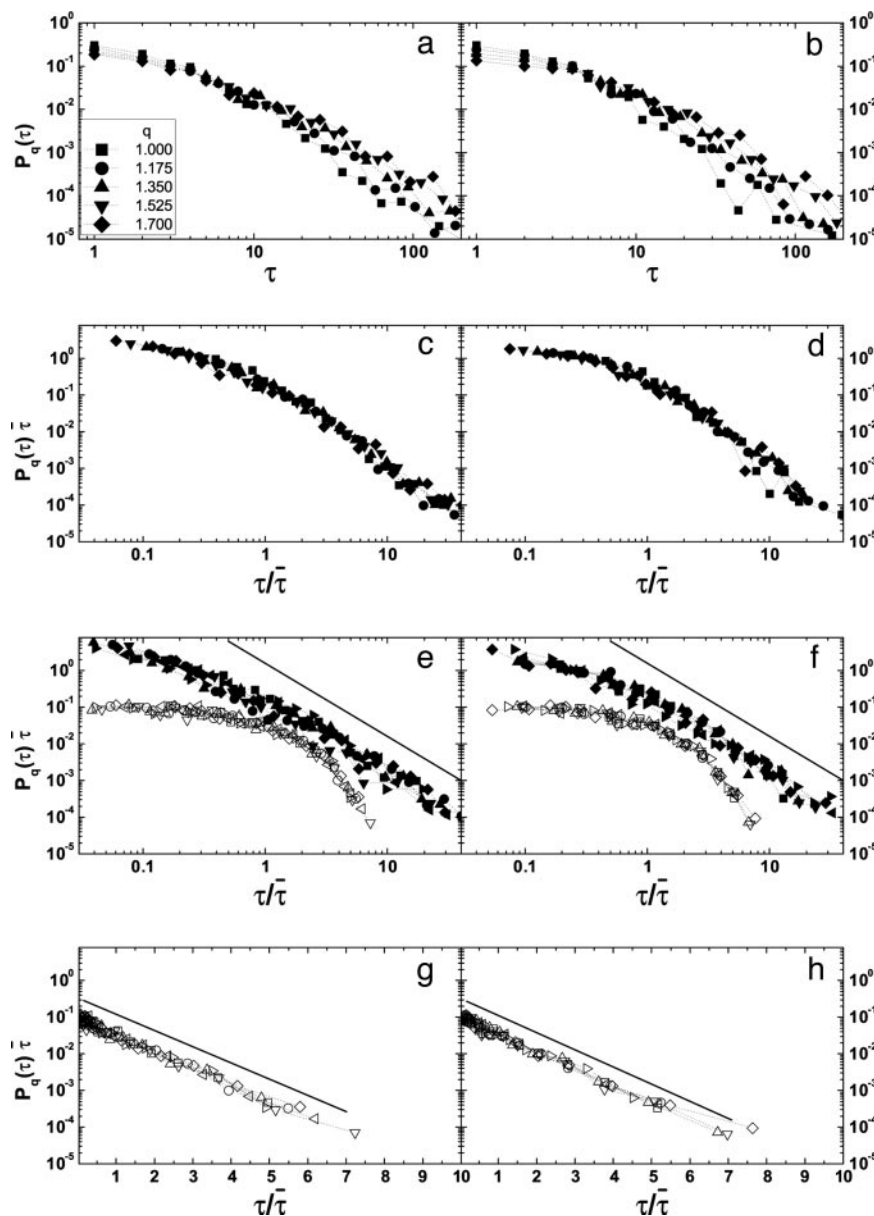


Fig. 2. Distribution of volatility return intervals. (a and b) Distribution function $P_q(\tau)$ of the return intervals τ of the volatility records of a typical stock, General Electric (GE) (a), and a typical currency exchange database, USD vs. SEK (b), for five threshold values q . (c and d) Scaled plots where $\tau P_q(\tau)$ is plotted vs. $\tau/\bar{\tau}$. (e and f) Scaled plots for seven stocks (e) and currencies (f). Stocks are S&P 500, \blacksquare ; IBM, \bullet ; DuPont, \blacktriangle ; AT&T, \blacktriangledown ; Kodak, \blacklozenge ; GE, \blacktriangleleft ; and Coca-Cola, \blacktriangleright . Currencies are USD vs. JPY, \blacksquare ; British pound vs. Swiss franc, \bullet ; USD vs. SEK, \blacktriangle ; Danish krone vs. Australian dollar, \blacktriangledown ; Danish krone vs. Norwegian krone, \blacklozenge ; USD vs. Canadian dollar, \blacktriangleleft ; and USD vs. South African dollar, \blacktriangleright . The straight lines are for demonstration and have a slope of -2 . The lower plots (open symbols) represent the scaled $P_q(\tau)\bar{\tau}$ vs. $\tau/\bar{\tau}$ after shuffling the volatility records, thereby removing the long-term correlations. The lower plots were divided by a factor of 10 [i.e., $P_q(\tau)\bar{\tau}/10$]. The open symbols represent the same stocks and the same currencies as the corresponding filled symbols in e and f. (g and h) To test whether after shuffling the volatility record $P_q(\tau)$ becomes a Poisson distribution, we plot on a semilogarithmic scale the shuffled stock (g) and currencies (h) data. The linear plot shows the exponential form of a Poisson distribution.

$\tau/\bar{\tau}$; the evident data collapse to a single curve is consistent with the scaling relation,

$$P_q(\tau) = \frac{1}{\bar{\tau}} f\left(\frac{\tau}{\bar{\tau}}\right). \quad [1]$$

We see that the scaling function $f(x)$ does not depend explicitly on q but only through $\bar{\tau} \equiv \bar{\tau}(q)$. Hence, if $P_q(\tau)$ is known for one value of q , Eq. 1 can make predictions for other values of q : in particular, for very large q (rare events), which are difficult to study because of the lack of data. The functional form of the scaled pdf appears to be quite similar for all seven stock records studied (Fig. 2e) and all seven currency data sets

studied (Fig. 2f). As suggested by the straight lines with slope -2 (Fig. 2e and f), $f(x)$ for $x \geq 1$ is consistent with a power law $f(x) \sim x^{-\gamma}$, where $\gamma \approx 2$ for both stock and currency data. For very large x values, the data are also consistent with the possibility that $f(x)$ is a stretched exponential (see also ref. 23). This result raises the possibility that there exists a “universal” scaling function for the return intervals of both stock and currency volatility data.

For uncorrelated records, we expect that the return intervals follow a Poisson distribution, yielding $\log f(x) \sim -x$. To test this expectation, we remove the long-term memory by shuffling the volatility records, and we obtain a simple exponential (Fig. 2g and h). The distinct difference between the distributions of the

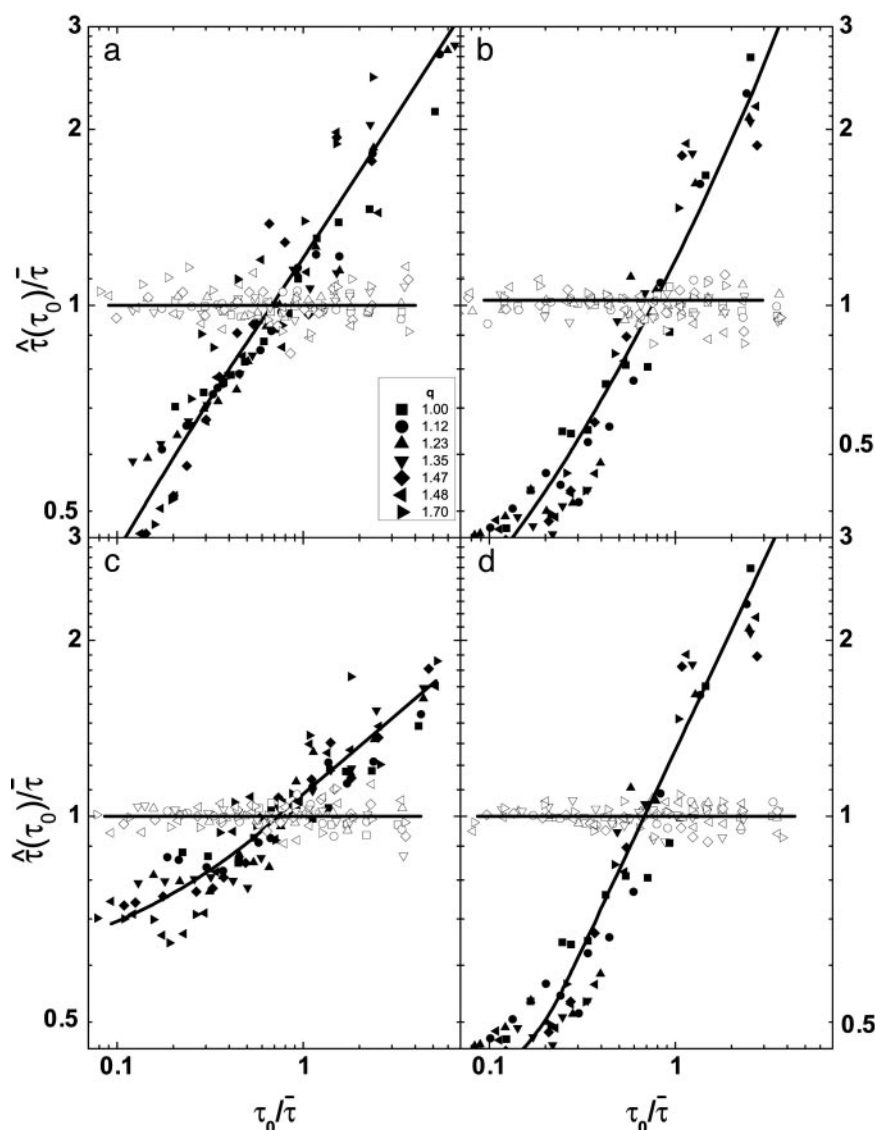


Fig. 5. Scaling and memory in mean return intervals. Shown is the mean conditional return interval $\hat{\tau}(\tau_0)$ divided by $\bar{\tau}$ vs. $\tau_0/\bar{\tau}$ of the volatility records for the daily S&P 500 (a), IBM stock (b), USD/JPY exchange rate (c), and USD/SEK exchange rate (d). For records without memory, we expect $\hat{\tau}(\tau_0)/\bar{\tau} = 1$, as supported by the open symbols obtained for volatility shuffled records.

interval τ_0 . In records without memory, $P_q(\tau|\tau_0)$ does not depend on τ_0 and is identical to $P_q(\tau)$. In records with long-term memory, we expect a pronounced dependence (23). To search for these kinds of memory effects in the (relatively) short data sets analyzed here, we study $P_q(\tau|\tau_0)$ not for a specific value of τ_0 but for a range of τ_0 values. For this purpose, we sort the full data set of N return intervals in increasing order and divide it into eight subsets, Q_1, Q_2, \dots, Q_8 , such that each subset contains one-eighth of the total number of return intervals. By this definition, the $N/8$ lowest return intervals are in Q_1 , whereas the largest $N/8$ intervals are in Q_8 . Fig. 4 *a* and *b* show, for the Standard and Poor's 500 Index (S&P 500) and for IBM stocks, $P_q(\tau|\tau_0)$ for τ_0 in Q_1 and Q_8 . The same quantities are studied in Fig. 4 *c* and *d* for the Japanese yen (JPY)/U.S. dollar (USD) exchange rate and the Swedish krona (SEK)/USD exchange rate. The results show that for τ_0 in Q_1 , the probability of finding τ below (or above) $\bar{\tau}$ is enhanced (or decreased) when compared with $P_q(\tau)$, whereas the opposite occurs for τ_0 in Q_8 . We note that, for the range of q studied, the data collapse onto a single scaling function for both Q_1 and Q_8 .

Results and Discussion

This memory effect in the conditional distribution function $P_q(\tau|\tau_0)$ leads to a pronounced memory effect in the mean conditional return interval $\hat{\tau}(\tau_0)$, which is the mean of those return intervals that immediately follow τ_0 . By definition, $\hat{\tau}(\tau_0)$ is the first moment of $P_q(\tau|\tau_0)$. Fig. 5 presents $\hat{\tau}(\tau_0)/\bar{\tau}$ as a function of $\tau_0/\bar{\tau}$ and shows clearly the effect of memory. Because small (or large) return intervals are more likely to be followed by small (or large) intervals, $\hat{\tau}(\tau_0)/\bar{\tau}$ is well below (or above) unity for $\tau_0/\bar{\tau}$ well below (or above) unity.

Thus, the statistics of the return intervals τ strongly depend on the preceding return interval τ_0 . We ask whether this memory is limited only to nearest-neighbor return intervals or whether there exists long-term memory in the return interval time series. To answer this question, we use detrended fluctuation analysis (DFA) to test for the presence of long-term correlations in τ (24). DFA calculates the rms fluctuation $F(\ell)$ of a time series within a window of ℓ days and determines the exponent α from the scaling relation $F(\ell) \sim \ell^\alpha$. When $\alpha > 1/2$, the time series is long-term correlated; when $\alpha = 1/2$, it is

